

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.

mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Apr 12 07:31:27 2000. MacDar + time 6 00 seconds

578.952 Million cell updates/sec

```
>US-09-276-268-12
Title:
Description:
Accession: US09276268 pep
Perfect Score: 133
Sequence: 1 MAPANIGLTPHVMIGAVI.
HDETFEFGVGLFESQNDVDPD 174
```

Scoring table: PAM 150

Gap 11

Searched: 82229 seqs, 29864866 residues

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Listing first 45 summaries

100

```
base: 8c3prot38
      1:swissprot
```

Statistics: Mean 44.875; Variance 85.350; scale 0.526

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Length	ID	Description	Ref. No.
1	1065	78.7	180	PITAVITIC SURFACE GLYCOP	9,676-159
2	1065	78.7	180	CELL SURFACE GLYCOP	9,838-98
3	129	9.5	1898	THY1 RABBIT	6,138-4
4	118	8.7	472	HYPOPHOSPHATE 54.3 K D P	6,138-4
5	113	8.4	453	THY1 CAEL	6,138-4
6	113	8.4	453	T-CELL SURFACE GLYCOP	3,436-03
7	114	8.4	453	CD4 CAEL	2,466-03
8	112	8.3	488	EXTRACELLULAR SURFACE RIBON	2,466-03
9	114	8.4	453	THY1 CAEL	6,138-4
10	112	8.3	488	SELECTING FACTOR U2AF	4,778-03
11	110	8.1	134	HYPOPHOSPHATE 57.1 K D P	6,138-03
12	110	8.1	134	CELL SURFACE GLYCOP	9,838-03
13	110	8.1	134	SMALL NUCLEAR RIBON	9,186-03
14	108	8.0	432	STYNOPTAGRAM II	1,766-02
15	108	8.0	432	STYNOPTAGRAM II	1,766-02
16	108	8.0	531	NON-MUSCLE CALDESMON	1,766-02
17	105	7.6	511	TRICHOTALIN	1,766-02
18	105	7.6	511	REV-INDUCED G PROTEIN-	2,426-02
19	105	7.6	511	HYPOPHOSPHATE 122.9 K D	4,586-02
20	105	7.8	1085	HYPOPHOSPHATE 122.9 K D	4,586-02
21	102	7.5	2468	MICROTUBULE-ASSOCIATED	4,586-02
22	104	7.6	1153	HYPOPHOSPHATE 29.0 K D P	6,386-02
23	103	7.5	591	SPERM PROTEIN N18	2,426-02
24	103	7.6	591	SPERM PROTEIN N18	2,426-02

ALIGNMENTS

	RESULT	I	STANDARD;	PRT:	180 AA.
AC	C11L ROMAN				
AC	PS3601:				
WT	01-OCT-1996 (Rel. 34, Created)				
WT	01-OCT-1996 (Rel. 34, Last sequence update)				
WT	15-JUL-1995 (Rel. 34, Last sequence update)				
DE	PUTATIVE SURFACE GLYCOPROTEIN C2IORF1 PRECURSOR (C2IORF3).				
OS	Homo sapiens (human).				
OS	C2IORF3				
NC	Proteins; Glycoproteins; Cranial; Vertebrate; Mammal;				
NC	Bacteria; Primates; Catarrhini; Reptile; Nonm.				
CC	[1]				
PP	SEQUENCE FROM N.A.				
PP	UNIVERSITY OF CALIFORNIA, SAN DIEGO				
EX	MEDLINE: 98231453				
EX	YASPO H., L., AALTONEN J., HOELLII-KUITONEN N., PELTONEN L., LEHRACH H.;				
A	"Cloning of a novel human putative type Ia integral membrane protein Genomics 49:133-136(1998).				
L					

-|- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN A CELL TRAFFICKING MECHANISM.

-|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See <http://www.isb-si.ch/annual/>)

```

C of send an email to license@sib-sib.ch).
C -----
C
C EMBL; Z50022; CAA90325.1; -.
C MIM: 603784.
C

```

TRANSMEMBRANE	GLYCOPROTEIN	SIGNAL	POTENTIAL
1	32	1	32
33	180	33	180

T	DOMAIN	33	96	EXTRACELLULAR (POTENTIAL).
T	TRANSMEN	97	117	POTENTIAL.
T	DOMAIN	118	180	CYTOPLASMIC (POTENTIAL).

	116	121	POLY-CYS.
T	DOMAIN		
T	CARBOHYD	45	POTENTIAL,
T	CARBOHYD	54	POTENTIAL,
T	SEQUENCE	50-55	60-65

Query Match 78.7%; Score 1055; DB 1; Length 180;
Best Local Similarity 75.1%; Pred. No. 9,678-199.

Matches 133; Conservative 25; Mismatches 16; Indels 3; Gaps 3;

[illegible]

RESULT	2	
ID	A33_HUMAN	
AC	000705	
STANDARD;		PRT: 319 AA.

Q37537
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
CELL SURFACE A33 ANTIGEN PRECURSOR

Homo sapiens (Human).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Primates; Catarrhini; Hominidae; Homo.
 [1]
 SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 DISCUSS

COON CARCINOMA. *Cell* 1974; 13: 1045.

HEATH J.K., KATZ S.J., JOHNSTONE C.M., CATMEL B., SIMPSON R.J., MORITZ R.L., TIGER G.P., JONES W.H., WHITEHEAD R.H., GROSEN L.G., SCOTT A.M., RUTTER G., COHEN L., WELLS S., OLD L.J., NICE E.C., BURGESS A.J.

The human A3 antigen is a transmembrane glycoprotein and a novel member of the immunoglobulin superfamily. *Proc. Natl. Acad. Sci. U.S.A.* 94:469-474 (1997).

POST-TRANSLATIONAL MODIFICATIONS.
MEDLINE, 97136159.
RITTER G., COHEN L.S., NICE E.C., CATIMEL B., BURGESS A.W., MORITZ R.G., JI H., HEATH S.J., WHITE S.J., KELLY S., OLD L.J., SIMPSON R.J.:
Characterization of posttranslational modifications of human A33 antigen and partially purified surface glycoprotein of human gastric carcinoma cells. *Int. J. Cancer*, 1988, 42, 103-110.
Blochem. Biophys. Res. Commun., 1986, 136, 469-473.

-1- FUNCTION: MAY PLAY A ROLE IN CELL-CELL COMMUNICATION AND SIGNALING
 -1- SOURCE: HUMAN COLON CANCER CELL LINE CACO-2 (1997)
 -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN
 -1- TISSUE SPECIFICITY: EXPRESSED IN NORMAL GASTROINTESTINAL
 PATHOLOGY AND IN 95% OF COLON CANCERS.
 -1- POSTTRANSLATIONAL MODIFICATION: GLYCOSYLATED, CONTAINS APPROXIMATELY 8 X OF N-LINKED
 CARBOHYDRATE
 -1- PTM: PALMITOYLATED
 -1- SIMILARITY: BELONGS TO THE IMMOBILINOGLOBULIN SUPERFAMILY. CONTAINS
 ONE C-2-LIKE AND ONE V-LIKE DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by other commercial entities requires a license agreement (see <http://www.ebi.ac.uk/ebis/ebis/ebis/>) or send an email to license@ebi.ac.uk.

EMUL: U71925; ANCS0957.1; -	
HSP: P06307; INED	
MIN: 602171; -	
PFAM: P06047; 19: 2	
PROSITE: P06047; 19: 2	
Transmembrane, Signal	
Antigen	
Lipoprotein; Palmitate; Glycoprotein;	
Signal	21
CHAIN	22 319
CELL SURFACE A13 ANTIGEN	
EXTRACELLULAR (POTENTIAL)	22 235

[illegible]

b	300 DYRQEQR	307
y	153 S-RHDEIR	159

RESULT	3	STANDARD;	PRT; 1998 AA.
C	FRANK ROMAN		
C	007181		
T	01-02-1994	(Rel. 30, Created)	
T	01-OCT-1994	(Rel. 30, Last sequence update)	
T	01-FEB-1995	(Rel. 31, Last annotation update)	
E	TRICHOHALIN		

THH OR TRYH OR THLH.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
SEQUENCE FROM H.A.
MEDLINE: 93280194.
LEE S.-C., KIM I.-G., MAREKOV L.N., O'KEEFE E.J., PASSEY D.A.

SEQUENCE OF 1731-1898 FROM N.A., AND CHARACTERIZATION.
MEDLINE: 91315897.

O'KEEFE E.J., HAMILTON E.H., LEE S.-C., STEINERT P.M.:
Keratinocyte: a structural protein of hair, tongue, nail, and
epidermis. *J. invest. Dermatol.* 101:655-715(1993).

IT MAY SERVE AS SCAFFOLD PROTEIN, TOGETHER WITH INOLUBLE FIBRIN, IN THE ORGANIZATION OF THE CELL ENVELOPE OR EVEN ANCHOR THE CELL ENVELOPE TO THE KIF NETWORK. IT MAY BE INVOLVED IN ITS OWN CALCIUM-DEPENDENT POSTMITOTIC PROCESSING DURING TERMINAL DIFFERENTIATION.

STUDENT: MONOMER (PROBABLE).

TISSUE SPECIFICITY: FOUND IN THE HARD KERATINIZING TISSUES SUCH AS THE INNER ROOT SHEATH (IRS) OF HAIR FOLLICLES AND DENTODULIA, AND IN THE FLATFOAM PAPILLAE OF DORSAL TONGUE EPITHELIUM.

1. DEVELOPMENTAL STAGE: EXPRESSED DURING LATE DIFFERENTIATION OF THE EPIDERMIS.

[illegible]

2

WALSH

(TM)

Release 3.1A John F. Collins, Biocompare Research, Inc.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Wsearch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Wed Apr 12 07:52:24 2000; MasPar time 7.02 Seconds
Tabular output not generated.
401,344 Million cell updates/sec

Title: >US-09-276-268-17
Description: 189632 seqs, 33686106 residues
Accession: 1 GTPASVLRNSVSDSLPPP.....LQGFSCVRLKMTWR 119
Sequence: 1 GTPASVLRNSVSDSLPPP.....LQGFSCVRLKMTWR 119

Scoring table: BAW 150
Gap 11

Searched: 189632 seqs, 33686106 residues

Post-processing: Minimum Match 0
Listing first 45 summaries

Databases: a:genbank35

Statistics: Mean 29.971; Variance 126.276; scale 0.237

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Length	DB	ID	Description	Pred. No.
1	763	85.5	98	1	W85614	Secreted protein clone	3.13e-63
2	420	47.1	56	1	V13130	Human secreted protein	1.90e-39
3	420	47.1	56	1	V13130	Human secreted protein	1.90e-39
4	93	10.4	482	1	R21409	MANH dehydrogenase sub	1.22e-01
5	88	9.9	167	1	W73235	Rabbit interferon-gamma	2.91e-01
6	88	9.9	167	1	W73235	Rabbit interferon-gamma	2.91e-01
7	88	9.9	167	1	W73235	Rabbit interferon-gamma	2.91e-01
8	86	9.6	273	1	R39284	Parainfluenza virus v	4.10e-01
9	85	9.5	272	1	W20657	H. pylori cytoplasmic	4.87e-01
10	85	9.5	300	1	W20958	H. pylori cytoplasmic	4.87e-01
11	83	9.3	341	1	W01456	E. coli O157:H7 p17 re	6.82e-01
12	83	9.3	463	1	W83215	O antigen fliptase inv	6.82e-01
13	83	9.3	463	1	W83215	O antigen fliptase inv	6.82e-01
14	82	9.2	333	1	R07923	Mouse acetylcholine re	6.07e-01
15	82	9.2	333	1	R07923	Mouse acetylcholine re	6.07e-01
16	81	9.1	131	1	W20762	H. pylori cell envelop	9.54e-01
17	81	9.1	131	1	W48222	A novel human leptin r	9.54e-01
18	81	9.1	131	1	W48222	A novel human leptin r	9.54e-01
19	81	9.1	334	1	W70954	Corn chloroplast nicot	9.54e-01
20	81	9.1	476	1	W70959	Lettuce infectious yel	9.54e-01
21	81	9.1	967	1	W93621	Human CD13/aminopeptid	9.54e-01
22	81	9.1	967	1	W93621	Human CD13/aminopeptid	9.54e-01
23	80	9.0	39	1	V13175	Braun-specific BS207	1.18e-02
24	80	9.0	39	1	V13175	Human secreted protein	1.18e-02

ALIGNMENTS

RESULT 1
ID W85614, standard; Protein; 98 AA.

DT 02-MAR-1999 (first entry)

DE Secreted protein clone fr473.2

KL clone; secreted protein; protein factor; cytokine; lymphokine;

KL clone; secreted protein; protein factor; cytokine; lymphokine;

KW tumor invasion; tumour suppression; immune boosting.

OS Homo sapiens.

PH Key

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

H. pylori membrane pro 1.13e-02
Human secreted protein 1.13e-02
S. pneumoniae cation t 1.13e-02
ORF 2 protein of PRSV 1.13e-02
Human secreted protein 1.13e-02
Human regulator of G-p 1.13e-02
Human acetylcholine re 1.13e-02
Nestin nist gene produc 1.13e-02
GJA-G-R subtype 1a. 1.13e-02
Amino acid sequence of 1.13e-02
Human hist acidu chan 1.13e-02
Helicobacter polyepit 1.13e-02
Staphylococcus aureus 1.13e-02
Condensing enzyme clon 1.13e-02
Human secreted protein 1.13e-02
Mutant Aspergillus ory 1.13e-02
Thermophilic DNA polym 1.13e-02
Alpha-DNA polymerase. 1.13e-02
Hatch a disease virus 1.37e-02

W85614, standard; Protein; 98 AA.

DT 02-MAR-1999 (first entry)

DE Secreted protein clone fr473.2

KL clone; secreted protein; protein factor; cytokine; lymphokine;

KL clone; secreted protein; protein factor; cytokine; lymphokine;

KW tumor invasion; tumour suppression; immune boosting.

OS Homo sapiens.

PH Key

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

FT Peptide

Query Match 85.5% Score 763; DB 1; Length 98;

Best Local Similarity 100.0%; Pred. No. 3.13e-63;

Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MASLCCGKCAAGVLSVANCYMLMGITFFNHNVLIEDVPTTFDENGQIYN 60

W85614, standard; Protein; 98 AA.

DT 02-MAR-1999 (first entry)

DE Secreted protein clone fr473.2

KL clone; secreted protein; protein factor; cytokine; lymphokine;

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

protein - protein database search, using Smith-Waterman algorithm

```

Title:
>US-09-276-268-20
Description:
(1-317) from US09276268.pep
Perfect Score:
2378
Sequence:
1 MRSGLWPLMGALVMTVGS.....KKAAGAGVTKSQAQAQAK 317

Scoring table:
PAM 150
Gap 11

```

```

Searched:      188963 seqs, 23686106 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database:      a-geneseq35
               l:geneseq35
Statistics:    Mean 33.327; Variance 130.318; scale 0.256

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the observed score. The observed
score is derived by analysis of the total score distribution and
is determined by the number of results predicted by chance to
have a score greater than or equal to the observed score.

```

SUMMARY

Ref.	No.	Score	Match	Length	ID	Description	Pred.	No.
2	1576	595	359	1	1008	Human follistatin-3 pr	6.1e-155	
3	789	595	359	1	97396	Human follistatin	6.1e-155	
4	789	332	317	1	99396	Human follistatin (hbs	1.2e-70	
5	789	330	304	1	82062	Rat EGF-binding prot	1.02e-69	
6	778	327	304	1	82063	Human EGF-binding prot	2.9e-69	
7	510	327	304	1	82064	Human EGF-binding prot	2.9e-69	
8	506	314	462	1	807624	Human polypeptide	2.02e-68	
9	506	313	462	1	81127	Elongation factor	6.58e-40	
10	428	280	400	1	80518	Prostate tumor induc	7.40e-39	
11	428	180	400	1	80518	plant elongation fac	9.9e-32	
12	393	165	293	1	805567	plant elongation fac	3.59e-28	
13	393	165	447	1	831584	61-1 clone polypep	4.5e-28	
14	388	163	448	1	830570	Translation elongate	1.17e-27	
15	388	163	448	1	831593	LEF1, protein	1.17e-27	
16	388	163	448	1	831595	LEF-1 clone A polypep	1.17e-27	
17	388	163	448	1	831595	plant elongation fac	1.17e-27	
18	188	79	664	1	803719	Protein elongation	4.2e-06	
19	170	71	380	1	817860	Follistatin growth fa	6.58e-06	
20	168	71	380	1	809406	Human GSRI-RK	3.5e-05	
21	143	60	303	1	805164	Human transforming	2.21e-03	
22	143	60	303	1	805164	Human transforming	2.21e-03	
23	140	59	374	1	807663	Human transforming	2.21e-03	

20

.....
 WILLOW (TM)

Release 3.1a John F. Collins, Bioinformatics Research Unit,
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MSPch_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Wed Apr 12 08:05:05 2000; MacPac Time 8.58 Seconds

Tabular output not generated. 478.775 Million cell updates/sec

Title: >US-09-276-268-20

Description: (1-317) from US09276268.pmp

Perfect Score: 2378

Sequence: 1 MESAQLPMLMALWTVGES.....KEAGAGVYKSAQAQAK 317

Scoring table: PAM 150

Searched: 131253 seqs, 1295647 residues

Post-processing: Minimum Match 04

Listing first 45 summaries

Database: a-issued

Statistics: 1:5A-COMB 2:5B-COMB 3:PC9-COMB 4:backfiles1

Mean 31.424; Variance 131.180; scale 0.236

Pred. No. is the number of results predicted by chance to have a
 man or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	1575	55	273	2	US-08-972-Sequence 2, Applicatio	1.03e-143
2	519	21.8	273	2	US-08-972-Sequence 2, Applicatio	1.41e-38
3	510	21.4	462	4	5253468-1 Patent No. 5253468-1	1.41e-38
4	500	21.4	462	2	US-08-371-Sequence 18, Applicatio	1.06e-37
5	500	21.4	462	2	US-08-371-Sequence 18, Applicatio	1.06e-37
6	488	20.5	398	2	US-08-299-Sequence 1, Applicatio	2.58e-37
7	483	20.3	397	2	US-08-371-Sequence 17, Applicatio	1.43e-35
8	453	19.0	421	2	US-08-972-Sequence 4, Applicatio	3.45e-35
9	453	19.0	421	2	US-08-972-Sequence 4, Applicatio	3.45e-35
10	188	7.9	1640	7	US-08-644-Sequence 30, Applicatio	1.97e-14
11	160	6.7	499	2	US-08-820-Sequence 46, Applicatio	1.91e-07
12	143	6.0	303	4	5340594-1 Patent No. 5340594-1	1.42e-05
13	143	6.0	303	4	5340594-1 Patent No. 5340594-1	1.42e-05
14	140	5.9	374	3	US-08-820-Sequence 25, Applicatio	2.53e-03
15	129	5.4	44	1	US-08-421-Sequence 2, Applicatio	2.53e-03
16	125	5.3	51	1	US-08-421-Sequence 2, Applicatio	2.53e-03
17	127	5.3	86	2	US-08-144-Sequence 1, Applicatio	4.95e-02
18	127	5.3	86	2	US-08-144-Sequence 1, Applicatio	4.95e-02
19	122	5.1	103	1	US-08-211-Sequence 3, Applicatio	2.98e-02
20	122	5.1	184	1	US-08-211-Sequence 3, Applicatio	2.98e-02
21	122	5.1	184	1	US-08-211-Sequence 3, Applicatio	2.98e-02
22	122	5.1	186	1	US-08-211-Sequence 15, Applicatio	7.57e-02
23	122	5.1	368	1	US-08-211-Sequence 37, Applicatio	7.57e-02

ALIGNMENTS

RESULT	ID	US-08-972-008-2	STANDARD;	PWT;	263 AA.
XX	XX	XXXXXX			
XX	XX	Sequence 2, Application US/08972008			
XX	XX	Sequence 2, Application US/08972008			
CC	CC	Patent No. 5342420-1			
CC	CC	GENERAL INFORMATION:			
CC	CC	TITLE OF INVENTION: N-3,3442420el Molecules of the Follistatin-Related			
CC	CC	INVENTOR: Douglas A			
CC	CC	NUMBER OF SEQUENCES: 5			
CC	CC	CORRESPONDENCE ADDRESS: 5			
CC	CC	STREET: 28 State Street			
CC	CC	CITY: Boston			
CC	CC	STATE: Massachusetts			
CC	CC	COUNTRY: USA			
CC	CC	ZIP: 02109			
CC	CC	COMPUTER READABLE FORM:			
CC	CC	MEDIUM TYPE: Floppy disk			
CC	CC	SOFTWARE: PC-DOS, MS-DOS			
CC	CC	OPERATING SYSTEM: PC-DOS, MS-DOS			
CC	CC	CURRENT APPLICATION DATA:			
CC	CC	FILE NAME: US/08972008			
CC	CC	CLASSIFICATION: 435			
CC	CC	ATTORNEY/AGENT INFORMATION:			
CC	CC	REGISTRATION NUMBER: 907			
CC	CC	REFERENCE/COCKET NUMBER: NHI-026			
CC	CC	TELECOMMUNICATION INFORMATION:			
CC	CC	TELEPHONE: (617) 742-4140			
CC	CC	TELEFAX: (617) 742-4140			
CC	CC	INFORMATION FOR SEQ ID NO: 2:			
CC	CC	SEQUENCE CHARACTERISTICS:			
CC	CC	TYPE: amino acid			
CC	CC	TOPOLOGY: linear			
CC	CC	MOLECULE TYPE: Protein			
CC	CC	SEQUENCE: 480 AA; 2603 AA; 26569 CH;			
CC	CC	Query Match			
CC	CC	Score 1575; DB 2: Length 263;			

